Chapter 13
Adaptive Fuzzy Inference Neural Network System for EEG and Stabilometry Signals Classification

Pari Jahankhani, Juan A. Lara, Aurora Pérez, and Juan P. Valente

Abstract. The focus of this chapter is to study feature extraction and pattern classification methods from two medical areas, Stabilometry and Electroencephalography (EEG).

Stabilometry is the branch of medicine responsible for examining balance in human beings. Balance and dizziness disorders are probably two of the most common illnesses that physicians have to deal with. In Stabilometry, the key nuggets of information in a time series signal are concentrated within definite time periods are known as events.

In this chapter, two feature extraction schemes have been developed to identify and characterise the events in Stabilometry and EEG signals. Based on these extracted features, an Adaptive Fuzzy Inference Neural network has been applied for classification of Stabilometry and EEG signals.

The model constructs its initial rules by a hybrid supervised/unsupervised clustering scheme while its final fuzzy rule base is optimised through competitive learning. A two-stage learning methodology is applied to this Neuro-Fuzzy structure, by incorporating gradient descent and recursive least squares estimations. The proposed modelling scheme is characterised by its high performance accuracy, high training speed and provides an efficient solution to the "curse of dimensionality" problem inherited in traditional neuro-fuzzy schemes.

In order to classify Stabilometric time series, a set of balance-related features have been extracted according to the expert's criteria.

The proposed Stabilometric medical diagnostic system is based on a method for generating reference models from a set of time series.

The experimental results validated the proposed methodology.

Pari Jahankhani
University of East London, School of Computing, Information Technology and Electronic, London E16 2RD, United Kingdom
e-mail: `pari@uel.ac.uk`

Juan A. Lara · Aurora Pérez · Juan P. Valente
Technical University of Madrid, School of Computer Science, Campus de Montegancedo, 28660, Boadilla del Monte, Madrid, Spain
e-mail: `j.lara.torralbo@upm.es,{aurora,jpvalente}@fi.upm.es`

N. Bessis, F. Xhafa (Eds.): Next Generation Data Technologies for CCI, SCI 352, pp. 329–355.
springerlink.com

1 Introduction and Background

The first medical diagnoses made by humans were based on what ancient physicians could observe with their eyes and ears. Most sophisticated diagnostic tools and techniques such as the thermometer for measure temperature and the stethoscope for measuring hear rate were not used until the end of 19^{th} century. In the 19^{th} century, diagnostic tools, including the microscope and X-ray helped provide hard data independent of subjective judgement and anecdote [1].

Number of medical diagnostic decision support (DS) systems based on computational intelligence methods has been developed for medical diagnosis. The medical diagnostic knowledge can be automatically derived from the description of cases solved in the past.

In medicine data mining is often used to complement and expand the work of the clinician and researcher by expanding knowledge rather than providing new knowledge as is the trend in other domains.

Data mining is the research area involving powerful processes and tools that allow an effective analysis and exploration of usually large amounts of data. Data mining techniques have found application in numerous different scientific fields with the aim of discovering previously unknown patterns and correlations.

Classifiers can play an intermediate role in multilevel filtering systems. In medical practice, the collection of patient data is often expensive, time consuming and harmful for the patients. Therefore it is necessary to have a classifier that is able to reliable diagnose with a small amount of data, also the process of determining the right subset of data may be time consuming as it is essentially a combinatorial problem.

The derived classifier can then be used either to assist the physician when diagnosing new patients in order to improve the diagnostic speed and accuracy.

Some medical diagnosis systems based on computational intelligence methods use expert systems (ESs) [2,3], fuzzy expert systems (FESs) [4,5,6], neural networks (NNs) [7,8,9,10] and genetic algorithms (GAs) [11]. Fuzzy Logic (FL) [12] is a "language", which uses syntax and local semantics where we can imprint any qualitative knowledge about the problem to be solved.

With the continuously growing demand for models of complex systems inherently associated with nonlinearity, high-order dynamics, time-varying behaviour, and imprecise measurements there is a need for a relevant modelling environment. Efficient modelling techniques should allow for a selection of pertinent variables and a formation of highly representative datasets. The models should be able to take advantage of the existing domain knowledge (such as a prior experience of human observers or operators) and augment it by available numeric data to form a coherent data knowledge modelling entity.

Fuzzy systems accept numeric inputs and convert these into linguistic values (represented by fuzzy numbers) that can be manipulated with linguistic IF-THEN rules and with fuzzy logic operations, such as fuzzy implication and composition rules of inference. However, at present there is no systematic procedure for the design of a fuzzy system. Usually the fuzzy rules are generated by converting human operators' experience into fuzzy linguistic form directly and by summarizing the system behaviour (sampled input-output pairs) of the operators. But designers find it difficult to obtain adequate fuzzy rules and membership functions because these are most likely to be influenced by the intuitiveness of the operators and the designers. Moreover, some information will be lost when human operators express their experience by linguistic rules. This results in a set of fuzzy rules which are usually not optimal. Thus a fuzzy system which is able to develop and improve its fuzzy rules and structure automatically on the basis of the monitoring of human controllers is highly desired.

During recent years, the fuzzy neural network approach has gained considerable interest for solving real world problems, including modelling and the control of highly complex systems, signal processing and pattern recognition [13]. Extensive experimentation has demonstrated that the class of feed-forward fuzzy neural networks exhibits a number of significant advantages compared to the neural network models [14]. First, the neural networks are global models where training is performed on the entire pattern range. On the contrary, owing to the partition of the input space, the fuzzy models perform a fuzzy blending of local models in space. As a result, faster convergence is achieved during learning for a specific task. Secondly, fuzzy neural networks are capable of incorporating both numerical data (quantitative information) and expert's knowledge (qualitative information) and describe them in the form of linguistic IF-THEN rules. In that respect, they provide a unified framework for integrating the computational parallelism and low-level learning of neural networks with the high-level reasoning of fuzzy systems. The above feature assists in the determination of the initial structure, also leading to models with fewer parameters compared to neural networks.

Neuro-fuzzy (NF) systems have been extensively used in pattern classification applications, including specialised medical ones [15]. Examples of NF systems as classifiers include schemes such as ANFIS [17, 18, 19], FSOM [20], Fuzzy-RBF [21], Fuzzy-ART [22]. The most famous example of Neuro-fuzzy network is the Adaptive Network based Fuzzy Inference System (ANFIS) developed by Jag [23] that implements a TS Fuzzy System [24, 25] in a multilayer architecture, and applies a mixture of both back-propagation and least mean squares procedure to train the system. In [26], a combination of fuzzy logic and neural network has been used to develop an adaptive control system for medical purpose.

1.1 EEG

An electroencephalogram (EEG) machine is a device used to create a picture of the electrical activity of the brain. It has been used for both medical diagnosis and neurobiological research. The essential components of an EEG machine include electrodes, amplifiers, a computer control module, and a display device.

EEG machines are used for a variety of purposes. In medicine, they are used to diagnose such things as seizure disorders, head injuries, and brain tumour.

EEGs are time-series signals with added noise. From such type of signals, many times a kind of information of the nature about these signals is required in real time in order to take crucial decisions [27, 28, 29, 30].

Electroencephalogram analysis is a very useful technique to investigate the activity of the central nervous system. It provides information related to the brain activity based on measurements of electrical recordings taken on the scalp of the subjects. Inference and studies about the subject's health and effective treatment of many diseases can be carried out by analysing the information obtained from the EEG.

EEG analysis has often been used to help doctors in Medicine and Information Technology to assist in their diagnostic procedures, especially whenever there are problems of different diagnosis in diseases. Methods from the domain of Intelligent Systems give the opportunity to formalise the medical knowledge and standardise various diagnostic procedures, in specific domains of Medicine and to store them in computer systems.

All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference. The data were digitised at 173.61 samples per second using 12 bit resolution. Band-pass filter settings were 0.53–40 Hz (12dB/oct).

Typical EEGs are depicted in Fig. 1.

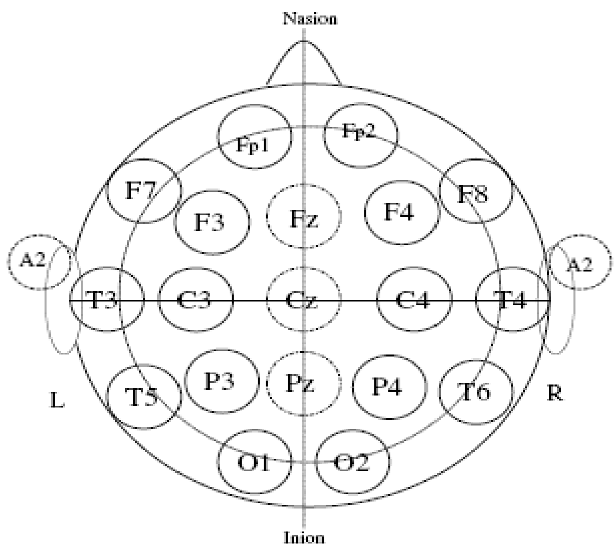

Fig. 1 The 10–20 international system of electrode placement c images of normal and abnormal cases.

The proposed EEG medical diagnostic system has been designed around the concept of a new neuro-fuzzy model.

The proposed diagnostic system, incorporating the proposed modelling scheme and also wavelet theory for the feature extraction section, achieved immensely satisfactory results on real test cases.

1.2 Stabilometry

Stabilometry is the branch of medicine responsible for examining balance in human beings. Balance and dizziness disorders are probably two of the most common illnesses that physicians have to deal with. Around 30% of population suffers from any kind of dizziness disorder before reaching the age of 65; for older people, this pathologic symptom occurs more frequently, and cause falling. The patient stands on a platform and completes a series of tests. These tests are designed to isolate the main sensorial, motor and biomechanical components that contribute to balance.

In order to examine balance, a device, called posturograph, is used to measure the balance-related functionalities. The patient stands on a platform and completes a series of tests, as shown in figure 2. These tests have been designed to isolate the main sensorial, motor and biomechanical components that contribute to balance. Emphasis has been given to the evaluation of the capacity for each individual components as well as the overall components capacity. The posturograph generates a time-series signal, where the main information normally is confined to events.

Fig. 2 Patient completing a test on a posturograph.

Initially, stabilometry was considered as a technique measuring only the balance of human beings under certain conditions [31]. Many researchers have studied the effect of closed eyes on balance [32] [33]. These works confirmed that the condition of having the eyes closed affects balance due to the fact that balance has a strong visual component.

Currently, stabilometry is also considered as a useful tool for diagnosing balance-related disorders like the Parkinson disease [34] or benign vertigo of childhood [35]. Regarding stabilometric data analysis, body sway parameters have been used for analysis balance-related functions [36], [37]. However, it appears that classic posturographic parameters, such as the measure of the sway of the centre of pressure [38] have failed in the detection of balance disorders [39]. The analysis of stabilometric time series using data mining techniques offers new possibilities. Recently, a new method has been developed for comparison of two stabilometric time-series [40]. This method calculates the level of similarity of two time-series and can be applied to compare either the balance of two patients or to study how the balance of one patient evolves with time. Stabilometry also plays an important role in the treatment of balance-related diseases. The NedSVE/IBV system has been utilised for the development of a new method that assists in the rehabilitation of patients who have lost their balance [41].

2 Objectives and Contributions

The proposed method builds reference models from a set of time series by means of the analysis of the events that they contain. This method is suitable for domains where the relevant information is concentrated in specific regions of the time series, known as events. The proposed method enables to define regions of interest according to the knowledge extracted from the domain experts, which is a plus compared with other methods addressing the time series as a whole without taking into account that certain regions can be irrelevant in the domain in question.

The technique developed throughout this article has been successfully applied to time series generated by Stabilometric devices that record the electrical activity of the brain. In this article, for classification of extracted features from Stabilometric events an Adaptive Fuzzy Inference Neural Network system (AFINN) has been applied.

3 Data Selection and Recording

3.1 EEG

We have used the publicly available data described in Andrzejak *et al.* [9]. The complete data set consists of five sets (denoted A–E) each containing 100 single-channel EEG segments. These segments were selected and cut out from continuous multi-channel EEG recordings after visual inspection for artefacts, e.g., due to muscle activity or eye movements. Sets A and B consisted of segments taken from surface EEG recordings that were carried out on five healthy volunteers. Volunteers were relaxed in an awake-state with eyes open (A) and eyes closed (B), respectively. Sets C, D, and E originated from EEG archive of pre-surgical diagnosis. : fig.3 shows examples of five different sets of EEG signals taken from different subjects.

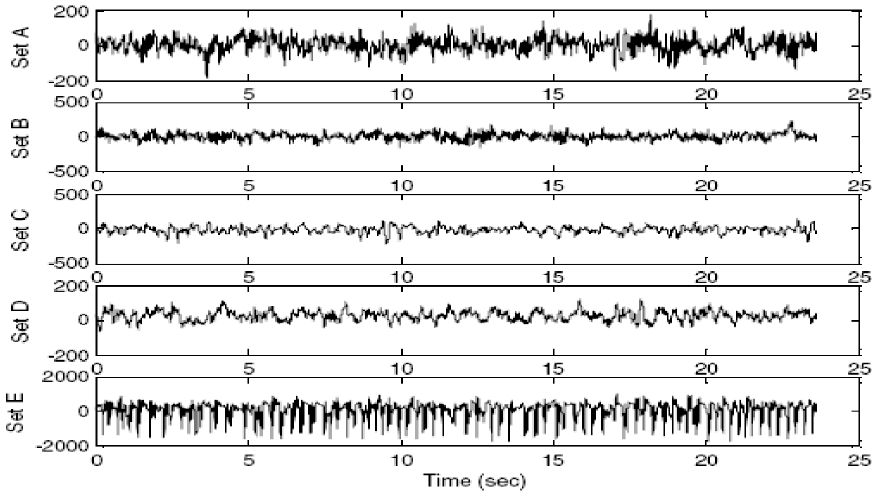

Fig. 3 Examples of five different sets of EEG signals taken from different subjects

3.1.1 Feature Extraction for EEG Signals

Numerous techniques from the theory of signal analysis have been used to obtain representations and extract the features of interest for classification purposes. Within this framework the signal is decomposed into sub-bands using fast wavelet transform algorithms.

The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the EEG signal in time and frequency. Table 1 presents frequencies corresponding to different levels of decomposition for Daubechies order-2 wavelet with a sampling frequency of 173.6 Hz. In order to further decrease the dimensionality of the extracted feature vectors, statistics over the set of the wavelet coefficients was used. The following statistical features were used to represent the time frequency distribution of the EEG signals:

- Maximum of the wavelet coefficients in each sub-band.
- Minimum of the wavelet coefficients in each sub-band.
- Mean of the wavelet coefficients in each sub-band
- Standard deviation of the wavelet coefficients in each sub-band

Table 1 Frequencies corresponding to difference levels of decomposition

Decomposed signal	Frequency range (Hz)
D1	43.4–86.8
D2	21.7–43.4
D3	10.8–21.7
D4	5.4–10.8
D5	2.7–5.4
A5	0–2.7

To reduce the volume of data, the sample (time points) was partitioned into 16 windows of 256 times points each. From these sub-samples, we performed the DWT and derived measures of dispersion statistics from these windows (each corresponding to approximately 1.5 seconds). The DWT was performed at 4 levels, and resulted in five sub-bands: d1-d4 and a4 (detail and approximation coefficients respectively). For each of these sub-bands, we extracted four measures of dispersion, yielding a total of 20 attributes per sample window. Since our classifiers use supervised learning, we must also provide the outputs, which was simply a class label.

3.2 Stabilometry

Throughout this research, a static Balance Master posturograph has been used. In a static posturograph, the platform on which the patient stands is static, i.e. does not move. The platform has four sensors, one at each of the four corners: right-front (RF), left-front (LF), right-rear (RR) and left-rear (LR). Each sensor records a datum every 10 milliseconds during the test. This datum is sent to the computer connected to the posturograph. The datum is the intensity of the pressure that the patient is exerting on that sensor. Data are recorded as multidimensional time-series.

The posturograph Balance Master can be used to run a wide range of tests according to a predefined protocol. This chapter has focused on the Unilateral Stance (UNI) test that is the most useful for domain experts (physicians) in terms of output information. UNI test aims to measure how well the patient is able to keep his or her balance when standing on one leg with either both eyes open or both eyes closed for 10 seconds. The UNI test generates time-series signals containing events, that is, regions of special interest for experts in the domain. Next section describes the possible events appearing in the time series of UNI test and the features used to characterise these events. Both the events and their features were determined according to the physicians' criteria. The following cases are the four different conditions of UNI test:

- Left leg with Open Eyes: The patient is asked to hold still with his or her left leg on the platform while his or her right leg has to be lifted.
- Right leg with Open Eyes: The patient is asked to hold still with his or her right leg on the platform while his or her left leg has to be lifted.
- Left leg with Closed Eyes: The patient is asked to hold still with his or her left leg on the platform while his or her right leg has to be lifted.
- Right leg with Closed Eyes: The patient is asked to hold still with his or her right leg on the platform while his or her left leg has to be lifted.

The current research has been carried out on time series from a set of healthy sportspeople, including both genders.

3.2.1 Model Generation Method

The model generation method presented here is suited for domains where only particular regions of the time series, known as events, contain relevant information for that domain while the remaining of the time series hardly provides any information. In order to deal with events, each event is characterized by a set of attributes.

The method proposed in this article receives a set of time series $S = \{S_1, S_2, ..., S_n\}$, each containing a particular number of events, and generates a reference model M that represents this set of time series. The model M is built on the basis of the most characteristic events. The most characteristic events of S are those events that appear in the highest number of timer series of S.

To find out whether a particular event in a time series S_i also appears in another time series S_j ($j \neq i$), the event has to be characterized with an attribute vector and compared with the other events of the other series. To speed up this process, all the events present in the time series are clustered, so similar events belong to the same cluster. On the one hand, the clustering process is useful to know the different groups of events. On the other hand, it facilitates the extraction of the most characteristic events. Once we have a set of clusters, the objective is to find those clusters containing events that appear in the highest number of time series, that is, characteristic events. Having located those groups with similar events, an exhaustive cluster analysis is run in order to extract the event representative of each of these groups. This will be described later (steps 5 to 9 of the algorithm). These extracted representative events are the characteristic events of S and will be part of the final model.

Let $S = \{S_1, S_2, ..., S_n\}$ be a set of n time series and m the typical number of events that appear in the time series of S. The algorithm for generating a reference model M representing the set S is as detailed below (with the purpose of making the algorithm more legible key decisions are justified at the end of the algorithm):

1. **Initialize the model.**
   M = Ø.
2. **Identify events.**
   Extract all the events E_v from the series of S and use an attribute vector to characterize each event. This vector covers what the expert considers to be the key features for each type of domain event. This step is domain dependent, as the event characterization will depend on the time series type. To extract the events, the time series is examined in search of regions that meet the conditions identifying each event type defined according to the knowledge extracted from the expert. Section 3.2.2 describes how the events are identified and characterised according to those conditions.
3. **Determine the typical number of events *m*.**
   m is the typical number of events in each time series of S. At the end of the algorithm it will be discussed how to determine this value.

4. Cluster events.
Cluster all the events extracted in step 2. Bottom-up hierarchical clustering techniques have been used. Taking into account that the proposal described here should be a general-purpose method and there is no a priori information for specifying the optimum number of clusters in each domain, bottom-up hierarchical clustering is a good option, as it is not necessary to specify the number of clusters k beforehand.

Repeat steps 5 to 9 *m* times
5. Get the most significant cluster C_k.
Determine which cluster C_k of all the clusters output in step 4 is the most significant. Cluster significance is measured using Equation (1).

$$\mathrm{SIGNF}(C_k) = \frac{\#\mathrm{TS}(C_k)}{n} \tag{1}$$

That is, cluster significance is given by the number of time series that have events in that cluster over the total number of time series n. Events that have already been examined (step 8 and 9) are not taken into account to calculate the numerator.
6. Extract the event E_c that best represents the cluster.
Extract the event that is most representative of the cluster C_k, that is, the event E_c that minimizes the distance to the other events in the cluster. Let S_j be the time series in which the event E_c was found.
7. Add the event E_c to the model.
$M = M \cup E_c$.
8. Mark event E_c as examined.
9. Mark the most similar events to E_c as examined.
From the cluster C_k obtain, for each time series $S_i \neq S_j$, the event E_p from S_i that is the most similar to the representative event (E_c) output in step 6. Each E_p will be represented in the model by the event E_c and therefore these E_p events will also be discarded in order not to be considered in later iterations.
10. Return *M* as a model of the set *S*.

The most significant clusters, that is, those clusters that contain events present in the highest number of time series were analysed to output the events that are part of the model. To do this, the process of identifying the most significant cluster is repeated m times, outputting a representative and marking as examined both this representative and similar events in each time series. With regard to the algorithm, note that:

a) The identification of events is domain dependent because the criteria to define events in each domain are required. The rest of the algorithm is domain independent and it can be applied to any domain without any change. Figure 2 shows the overall structure of the proposed method that receives a set of time series S and generates a model M that represents it.

b) After the representative event of the most significant cluster has been output, it should not be taken into account again for the next iteration, and it is marked as an already examined event.
c) A cluster may contain not just one but several events from each time series. For this reason, even if a cluster is selected as the most significant, the cluster in question is not omitted in later iterations. The events already processed are marked as examined and will not be taken into account in future iterations.

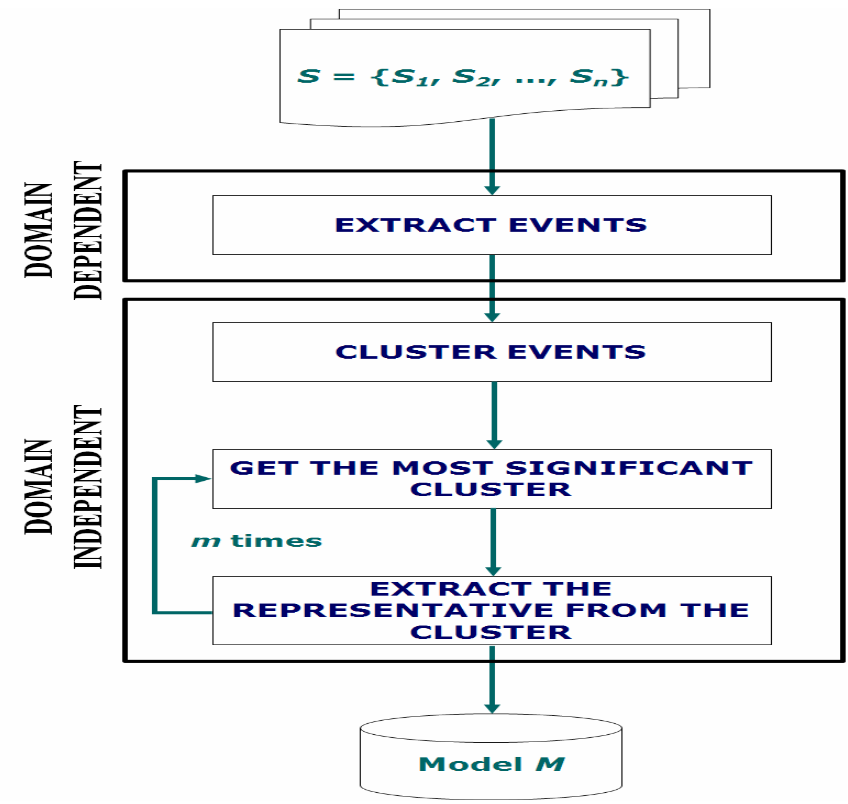

Fig. 4 Overall structure of the proposed method.

Another important issue is the number of events making up the model. In this case, we have chosen the mode (m) of the number of events of the time series of S. This decision is based on the fact that if the original time series have a typical number of events m, it makes sense for the model that represents them to have the same number of events m. The typical number of events in the time series of S

may not be unimodally distributed. This could happen especially if there are not many time series in the set S. For non-unimodal distributions, we have opted to take the integer value closest to the mean of the number of events.

A last point to be considered is the distance between events that has been used in the algorithm for clustering, representative event selection and discarding similar events. The *city block* distance is used. Given two vectors, the city block distance calculates the sum of the absolute value of the difference of each of the coordinates of the above vectors:

$$d(x, y) = \sum_{i=1}^{p} |x_i - y_i| \tag{2}$$

In Equation (2), x and y are the vectors (that is, the event descriptors) for comparison and p is the number of coordinates (dimension). Other distance measures have been studied, but the city block distance was finally chosen. The main reason for this choice is that the clustering algorithm uses the mean distance per attribute as the threshold for determining whether or not two elements are similar enough to belong to the same cluster. This mean distance per attribute is obtained simply by dividing the total city block distance $d(x,y)$ by the number of attributes p. The use of the city block distance then saves time as it obviates additional transformations that would make the clustering process more complex to develop and more computationally intensive.

Figure 5 shows an example of the application of the proposed method to a set $S = \{S_1, S_2, S_3, S_4\}$ of 4 time series (n=4). In this case, S_1 contains 2 events (E_{11} and E_{12}, S_2 contains 2 events (E_{21} and E_{22}), S_3 contains 3 events (E_{31}, E_{32} and E_{33}) and finally S_4 contains 2 events (E_{41} and E_{42}). Therefore, the typical number of events is 2 (m=2). Once the events are extracted, they are clustered into three different clusters (C_1, C_2 and C_3). Then, the most significant cluster is obtained. To do that, it is necessary to calculate the significance of each cluster according to Equation (EQ). In this case, cluster C_1 have events present in 3 out of the 4 time series, cluster C_2 have events that appear in 1 out of the 4 time series and cluster C_3 have events present in 4 out of the 4 time series of S. Then, the significance of C_1 is $SIGNF(C_1) = \frac{3}{4} = 0.75$, the significance of C_2 is $SIGNF(C_2) = \frac{1}{4} = 0.25$ and the significance of C_3 is $SIGNF(C_3) = \frac{4}{4} = 1$. Therefore, the most significant cluster is C_3. In the next step, the event E_{12} is extracted as the representative event of the cluster C_3 because E_{12} is the event in C_3 that minimizes the distance to the other events in that cluster. Thus, the event E_{12} is a characteristic event of S and will be part of the final model M. This process has to be repeated twice (because m=2) to build the final model that consists of the events E_{12} and E_{32}.

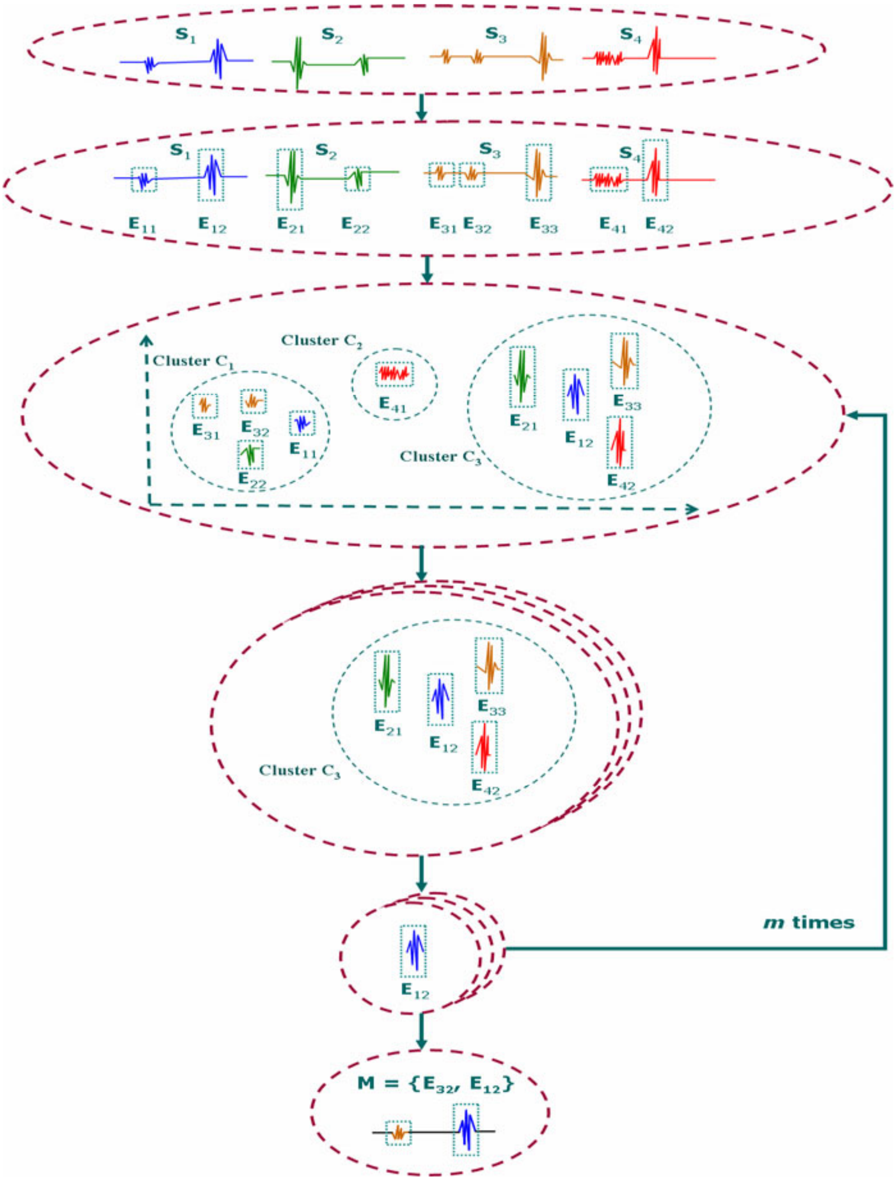

Fig. 5 Example of the application of the proposed method

3.2.2 Feature Extraction of Stabilometric Events

There are several stabilometric tests that can be carried out on the posturograph. In this chapter we have focused on the Unilateral Stance (UNI) test that is the most useful for the experts on the domain in terms of output information. The UNI test aims to measure how well the patient is able to keep his or her balance when standing on one leg with both eyes either open or shut for 10 seconds (see figure 5).

The ideal situation for the UNI test would be for the patient not to wobble at all but to keep a steady stance throughout the test. According to the knowledge extracted from the expert physicians, the interesting events in this test occur when the patient becomes unsteady, loses balance and puts the lifted leg down on the platform. This type of event is known in the domain as a fall. The features characterising the falls are as follows:

- **Duration:** It is the amount of time between the moment when the patient starts to lose balance and the moment when he or she is stable again, after falling.
- **Intensity:** It is the strength that the patient exerts on the platform when he or she falls down onto it.
- **Test time at which the event occurs:** It is the timestamp when the fall starts.

When there is a fall, the respective sensors for the lifted leg will register the pressure increase. Figure 6 shows the time-series of a patient who has taken the UNI test. The curves at the top of the figure are the values recorded by the RR and RF sensors, that is, the right leg sensors, the leg the patient was standing on. The curves at the bottom of the figure are the values recorded by sensors LR and LF, that is, the left-leg sensors, the leg that should be lifted. The pressure peaks generated when there is a fall event are highlighted. Figure 6 shows that the LR and LF time series are almost static (with a small variation margin) throughout, except when there is a fall. Therefore, it could be possible to define a stability value for the pressure exerted on the respective sensors for the above time series. This stability value, which in the case of this example would be around 20, could be a statistical measure like the mode. Every time there is a fall, there is a particular time instant where the LF and LR sensors record a local maximum and the RR and RF sensors record a local minimum. This point is more or less midway through the fall.

To identify the fall events and determine their features, the proposed method calculates the mode of the time-series related to the leg that must be lifted (bottom of figure 6). This value represents the balance value as shown in figure 7. The method identifies points where there is a local maximum whose distance to the balance value is higher than a certain threshold (∂). That distance is precisely the intensity of the fall. The duration of the fall is then calculated by analysing the two intersections between the time series and the balance value line. The timestamp at which the event starts is the first intersection between the time series and the balance value line.

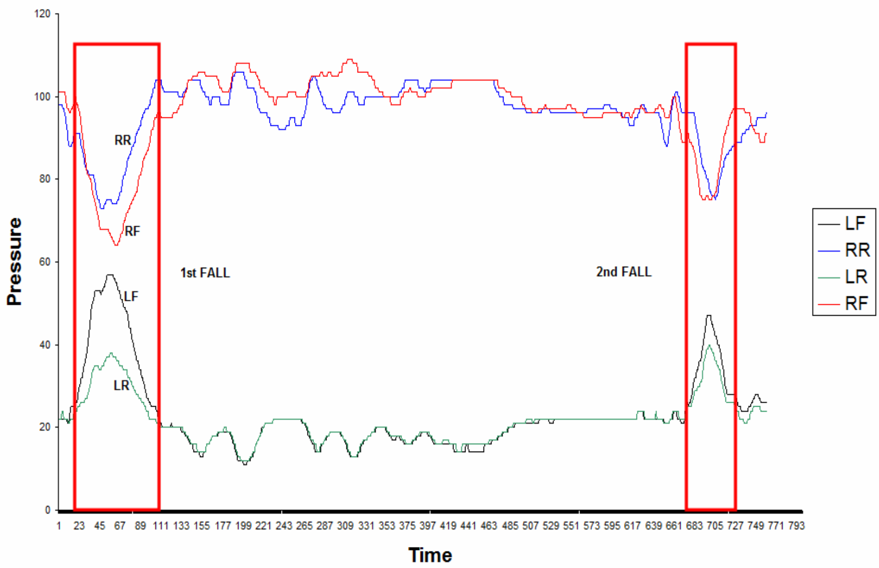

Fig. 6 UNI test time series, highlighting two events (falls)

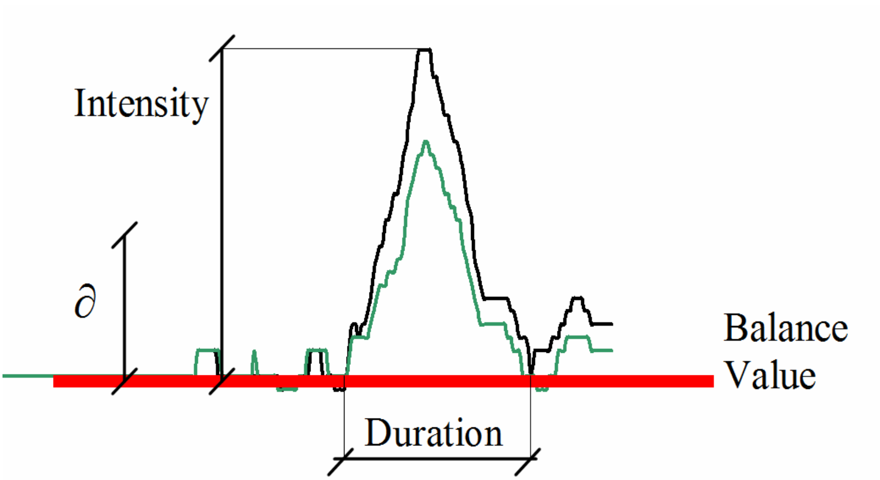

Fig. 7 Fall event taken from a stabilometric time series

4 Architecture of AFINN

There are many different combinations of fuzzy logic systems and neural networks.

This thesis proposes a connectionist model of fuzzy system in the form of feed-forward multi-layer networks, which can be trained using an iterative algorithm. This kind of neuro-fuzzy system employs a perceptron-like structure and a hybrid supervised learning procedure of neural networks for fuzzy inference system with rule base and fuzzy reasoning. The most important problem in fuzzy systems is to find fuzzy rules. Some methods can generate fuzzy rules from input-output pairs [42].

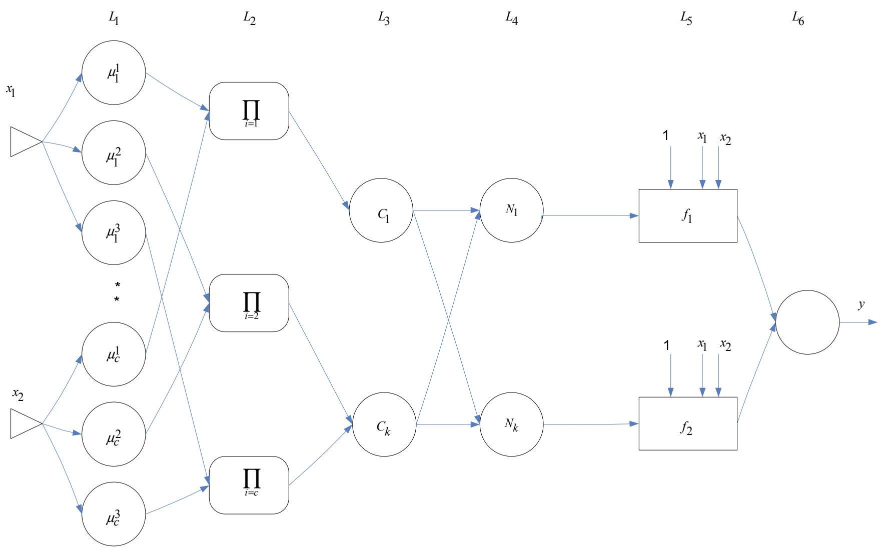

Fig. 8 Structure of AFINN system

The architecture of the proposed neuro-fuzzy network is shown in Figure 8 which consists of five layers. The first two layers L1 and L2, premise part correspond to the fuzzification section IF part of fuzzy rules whereas layers L4 and L5 the consequence part contain information about the THEN part of these rules and perform the defuzzification task. In Layer L3 a mapping between the rule layer and the output layer is performed by a competitive learning process. The local linear systems at L4 are associated with each term of layer L3 rather than that of rule base layer L2. Thus the size of the required matrices for least-squares estimation is considered to be much smaller.

The clustering algorithm which is applied in this framework at Layer L2 consists of two stages. In the first stage the method similar to the LVQ algorithm generates crisp c-partitions of the data set. The number of clusters c and the cluster centres $v_i, i = 1, \ldots\ldots, c$ obtained from this stage are used by the FCM algorithm in the second stage.

- The first stage clustering algorithm determines the number of clusters by dividing the learning data into these crisp clusters and calculates the cluster centres which are the initial values of the fuzzy cluster centres derived from the second stage algorithm. Let $X = [x_1, \ldots\ldots, x_n] \in R^{np}$ be learning data. The first cluster is created starting with the first data vector from X and the initial value of the cluster centre is taking as a value of this data vector. Then other data vectors are included in the cluster but only the ones that satisfy the following condition

$$\left\| x_k - v_i \right\| < D \tag{3}$$

Where $x_k \in X, k = 1,....,n$ and $v_i, i = 1,........,c$ are cluster centres, $[v_1,....,v_n] \in R^{cp}$, the constant value D is fixed at the beginning of the algorithm. Cluster centres v_i are modified for each cluster (i.e., $i = 1,.....,c$) according to the following equation

$$v_i(t+1) = v_i(t) + a_t(x_k - v_i(t)) \tag{4}$$

Where $t = 0,1,2,......$ denotes the number of iterations, $a_i \in [0,1]$ is the learning rate and decreases during the performance of the algorithm (depending on the number of elements in the cluster). Recursion of Eq. 4, originates from the LVQ algorithm. As a result of the performance of this algorithm we get the number of clusters c that we have divided the data set into, and thus we know the values of the cluster centres $v_i, i = 1,........,c$ which we can use as initial values for the second stage clustering algorithm.

- In the second stage the fuzzy c-means algorithm has been used. FCM is a constrained optimisation procedure which minimises the weighted within-groups sum of squared errors objective functions J_m with respect to the fuzzy membership's u_{ik} cluster centres v_i, given the training data $x_i, \;\; i = 1,....,c \quad k = 1,.....,n$

$$\min_{(U,V)} \{ J_m(U,V;X) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m \left\| x_k - v_i \right\|_A^2 \} \tag{5}$$

Where $U = [u_{ik}]_{c \times n}$ and u_{ik} 's satisfy the following conditions:

$$\begin{aligned} & 0 \le u_{ik} \le 1 && \forall_{i,k} \\ & 0 < \sum_{k=1}^{n} u_{ik} < n && \forall_i \\ & \sum_{i=1}^{c} u_{ik} = 1 && \forall_k \end{aligned} \tag{6}$$

The matrix $U \in R^{cn}$ is called a fuzzy c-partition of X, where $X = [x_{1,......,x}$ X, where $X = [x_1,........,x_n] \in R^{np}$ is a data set.

The distance between x_k and v_i is the Euclidean norm when the distance matrix A is chosen as the identity matrix I.

The centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{7}$$

The optimal values of u_{ik} and v_i which create the optimal fuzzy c-partition of the data set into c clusters are calculated as follows:

$$u_{ik} = \left[\sum_{j=1}^{c} \left(\frac{\|x_k - v_i\|_A}{\|x_k - v_j\|_A} \right)^{2/(m-1)} \right]^{-1} \qquad \forall_{i,k} \tag{8}$$

Formulas (7) and (8) are repeated until J_m no longer decreases.

The number of clusters c and the initial values of cluster centres v_i come from the first stage clustering algorithm. The most important problem in fuzzy systems is to find fuzzy rules. In this chapter the fuzzy rule base is derived using results obtained from the clustering algorithm described in this section.

In our proposed system L3 (Layer 3) is an additional layer of output partitions, each of which is associated with a local number of rule nodes which will be reduced and each node in Layer 3 will only be connected to one node in Layer 2. Nodes in this layer represent the partitions of the output variables, link at this layer form as consequences of the rules confusing for me.

The main rational underlying the work described in this chapter, which represents the core of the thesis is the development of a new neuro-fuzzy network. We will consider an Adaptive Fuzzy Inference Neural Network system (AFINN) which is made up of Gaussian-membership functions associated with local linear systems. The proposed fuzzy logic system is based on the Sugeno-type modified with the introduction of an additional layer of output partitions. Unlike the ANFIS system, in which the number of local linear systems is the same as that of the number of rules, AFINN provides a means of controlling the growth of the number of local linear systems when the order of the system under consideration increases, so that least-squares estimation can be applied without performance degradation. A clustering algorithm is applied for the sample data in order to organise feature vectors into clusters such that points within a cluster are closer to each other than vectors belonging to different clusters. Then fuzzy rules will be created using results obtained from this algorithm. Unlike Sugeno's method [43], the fuzzy implication of the fuzzy system is based on fuzzy partitions of the input space directly rather than fuzzy partitions of each dimension of the input space. Thus the membership functions considered in the proposed system are multidimensional membership functions. In this sense, there is a similarity with the construction of Gaussian centres in Radial Basis Function networks (RBFN). Since the input space is considered to be partitioned instead of each dimension of the input space, the number of rules can be small and hence the number of local linear systems is also small. In addition, a competitive learning technique is applied to locate space partitions according to the clustering of the fuzzy rules at the beginning of training. The proposed methodology is implemented and its performance evaluated against Multilayer Perceptron (MLP).

Essentially, we can say that the interpretability of fuzzy model depends on two main aspects: the complexity of the fuzzy rule base (i.e. the number of rules) and the readability of the fuzzy sets used in the fuzzy rules.

The number of rules is a crucial condition to obtain linguistically interpretable fuzzy models. Therefore the problem of finding a balance between the complexity (rule base size) and the accuracy of the fuzzy model is of considerable importance.

4.1 Training Procedure for AFINN

In the tuning phase, emphasis is given to the nature of AFINN scheme itself. Two different sets of parameters exist and need to be tuned. These include the nonlinear premise parameters in the fuzzification part and the linear consequent parameters in the defuzzification part. A hybrid learning approach thus has been adopted for the AFINN scheme. The network can be considered as a cascade of nonlinear systems and linear systems. In this phase, the error back-propagation is applied to tune the premise parameters of the membership functions and recursive least squares estimation is applied to find the consequence parameters of local linear systems.

The hybrid algorithm is composed of a forward pass and a backward pass. The least squares method is used to optimize the consequent parameters with the premise parameters fixed. Once the optimal consequent parameters are found, the backward pass starts immediately.

5 Evaluation

5.1 Evaluation of the Model Generation Method

The chapter focused on the UNI test data. Thorough the evaluation process, we used time series taken from a total of 30 top-competition sportspeople, divided into two groups. The first group was composed of 15 professional basketball players, whereas the second was made up of 15 young elite skaters. Thirty is a reasonable number of patients, taking into account that top-competition athletes do not abound and the stabilometric tests are quite complex (a single patient check-up takes up 2-3 Mb).

The ultimate aim of the evaluation is to measure how good the model generation method is. Two models from each of the above groups of sportspeople have been created. The first model ($M_{basketball}$) was created from a training set composed of 10 of the 15 basketball players. The other 5 players constituted the test set. The second model ($M_{skating}$) was generated from a training set composed of 10 of the 15 skaters. The other 5 skaters were used as test set. The sportspeople in the test set were chosen at random from all the sportspeople in each group.

Once the models have been created, they have been evaluated by checking whether the $M_{basketball}$ model properly represents the group of professional basketball players and whether the $M_{skating}$ model is representative of the group of elite skaters. To do that, we have compared each of the ten individuals in the test group against each of the two created models, making use of the Stabilometric time series comparison method described in [37]. This process was repeated five times changing the training set and the test set. The final results show that the 90% of sportspeople across all the experiments were successfully classified. It demonstrates the ability of the proposed method to generate representative reference models in the field of Stabilometry.

5.2 Evaluation of AFINN

Classification problems, also referred to as pattern recognition tasks, arise in a variety of engineering and scientific disciplines such as biology, psychology, medicine and artificial intelligence.

To illustrate the applicability of the proposed methodology to classification tasks, four classification datasets have been considered: the EEG data set, Cancer dataset, the Iris dataset and the MONK's dataset.

To assess the prediction quality of the resulting model on the basis of the available data, we have to estimate the unknown prediction risk; the estimate of the true generalization error of a model based of a finite set of observations can be obtained via data re-sampling techniques such as holdout, cross-validation and bootstrap. The basic idea is to build a model, from one part of the available data, and then use that model to predict the rest of the data. A 5-fold cross validation was performed, which involves splitting the whole data set into 5 equally sized parts and using each part as a test set for the model extracted from the remaining data. The estimate of the generalisation error is computed as the average of classification errors provided by models extracted from the 5 partitions.

Split the dataset into two groups: The Training set is used to train the classifier and testing set to estimate the error of the trained classifier.

The classification methodology proposed in this thesis and described in this chapter has been tested on several benchmark examples for classifications.

The results of these tests can be found in many papers published on international journal and conference proceedings which contributed to the work of this chapter [44, 45, 46 and 47].

These benchmark examples cover a wide range of applications.

All experiments have been tested using the MATLAB software package.

The training data set was used to train the AFINN model, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained AFINN model for classification

6 Discussion of Results

Automated diagnostic systems aim to enhance the ability to detect pathological structures in medical examinations and to support evaluation of pathological findings during the diagnostic procedure. Most techniques developed for automated Stabilometric data analysis have focused on the study of the centre of pressure of the patient. However, balance-related events (falls and imbalances) contain useful information for the physicians. In this research, the proposed AFINN network has been implemented for Stabilometric time-series classification, employing the most significant features of the events contained in UNI time series. As explained in section 2, the UNI test consists of four different conditions. This case study focused on the second trial of Left leg with Closed Eyes condition. The first and third trials have not been considered because the first trial contains noise and during the third trial the patient has already learnt how to be stable. In this chapter, 56 Stabilometric time series with 1000 timestamps have been used. The data set was divided into two classes according to the gender of patients: MALE and FEMALE. 38 out of the 56 time series belong to male patients while 18 belong to female patients. 5 balance-related features were extracted from each time series.

In our experiments, the training data set was used to train the AFINN model, whereas the testing data set was used to verify the accuracy and effectiveness of the trained AFINN model for classification for the 2 classes of Stabilometric time series. The proposed scheme has high classification accuracy with within 5 epochs. The results of the proposed classifier, using 10 different training sets for Stabilometry are illustrated at Table 2.

Table 2 AFINN performance for Stabilometric time series

System	Rules or number of Nodes	Epoch	Class 1 (Female)	Class 2 (Male)
AFFIN	7/4	5	95%	94.3%

The clustering Fuzzification part resulted in 7 rules, while after the competitive layer, the rules were reduced to 4, which resulted in fewer consequent parameters at the Defuzzification layer.

AFFIN performance for A, B, C, D and E classes illustrated in table 3.

Table 3 AFINN performance for A, B, C, D, and E classes

Data sets	No. rules	Class A	Class B	Class C	Class D	Class E
1	9/5	96.88	99.4	98.44	98.44	97.5
2	9/5	99.4	98.4	99.1	99.1	99.4
3	6/4	98.75	99.4	99.4	99.1	98.75
4	9/5	99.4	99.1	96.88	98.75	98.75
5	7/5	98.44	98.44	98.75	99.1	96.88
Average	7/4	98.6	99.1	98.5	98.9	98.26

The training data set was used to train the AFINN model, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained AFINN model for classification of the original five classes of EEG signals.

7 Conclusions

The historical development of machine learning and its applications in medical diagnosis shows that from simple and straightforward to use algorithms, systems and methodologies have emerged that enable advanced and sophisticated data analysis.

Fuzzy set theory plays an important role in dealing with uncertainty when making decisions in medical applications. Using fuzzy logic has enabled us to use the uncertainty in the classifier design and consequently to increase the credibility of the system output.

With the recent advancement of computational technologies along with the technologies in health care and biomedical field, the research community in medical intelligent systems and in machine learning is facing new challenges and opportunities to contribute substantially to clinical medicine.

Several hybrid systems combining different soft computing paradigms such as neural networks, fuzzy systems and genetic algorithms, have been proposed as a technique ideal for predictive modelling or classification. In particular considerable work has been done to integrate the excellent learning capability of neural networks with the representation power of fuzzy inference systems. This results in neuro-fuzzy modelling approaches, which combine the benefits of these two powerful paradigms into a single case and provide a powerful framework to extract fuzzy rules from numerical data. The aim of using a neuro-fuzzy system is to find, through learning from data, a fuzzy model that represents the process underlying the data.

A neuro-fuzzy methodology for modelling is presented as a neural network implementation of the new fuzzy system. We have studied a two stage clustering algorithm to determine the rules, number of fuzzy sets, and the initial values of the parameters i.e. the centres and widths of the fuzzy membership functions. Unlike ANFIS in which the number of local linear systems is the same as that of the rules, the proposed system provides a means of controlling the growth of the number of local linear systems when the order of the system under consideration increases so that least square estimation can be applied. The proposed network was trained and tested with the extracted features using a discrete wavelet transform of the EEG signals, and then Principal Component Analysis (PCA) was used to reduce the data dimensionality.

The simulation results reveal an almost perfect performance compared to the classic MLP neural network.

The proposed network was trained and tested with the extracted features using a statistical method for the identifications and characterisation of events in Stabilometric time series.

We have developed a method to generate reference models from a set of time series by matching up the events that they contain. This method is suitable for domains where the key information is concentrated in specific regions of the series, called events, and the remaining regions are irrelevant. The proposed method enables to define regions of interest according to the knowledge extracted from the domain experts, which is a plus compared with other methods addressing the time series as a whole without taking into account that certain regions can be irrelevant in the domain in question.

The method was evaluated on stabilometric time series, obtaining very satisfactory results, especially as regards the representativeness of the reference models generated by the proposed method. The results confirm the generality of the model generation method.

The proposed AFINN scheme characterised by its high performance accuracy, high training speed provides an efficient solution to the "curse of dimensionality" problem inherited in the classical NF scheme.

References

[1] Baxt, W.G.: Application of Ariticial Neural Networks to Clinical Medicine. Lancet 346, 1135–1138 (1995)
[2] Jang, J.-S.R.: ANFIS: Adaptive Network Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics (1993)
[3] Jang, J.-S.R.: Neuro-fuzzy Modelling and Control. The Proc. of the IEEE (1995)
[4] Sugeno, M., Kang, G.T.: Structure identification of fuzzy model. Fuzzy Sets and systems 28, 15–33 (1988)
[5] Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modelling and control. IEEE Trans. On systems, Man, and Cybernetics 15, 116–132 (1985)
[6] Kovalerchuk, B., Vityaev, E., Ruiz, J.F.: Consistent Knowledge Discovery in Medical; Diagnosis. IEEE Engineering in Medicine and Biology Magazine 19(4), 26–37 (2000)
[7] Wiegerinck, W., Kappen, H., ter Braak, E., Nijman, M., Neijt, J.: Approximate inference for medical diagnosis. Pattern Recognition Letters 20(11-13), 1231–1239 (1999)
[8] Kovalerchuk, B., Vityaev, E., Ruiz, J.F.: Consistent Knowledge Discovery in Medical; Diagnosis. IEEE Engineering in Medicine and Biology Magazine 19(4), 26–37 (2000)

[9] Walter, D., Mohan, C.: ClaDia: a fuzzy classifier system for disease diagnosis. In: Proceedings of the 2000 Congress on Evolutionary Computation, pp. 1429–1435 (2000)
[10] Zahan, S.: A fuzzy approach to computer-assisted myocardial Ischemia diagnosis. Artificial Intelligence in Medicine 21(1-2), 271–275 (2001)
[11] Pena-Reyes, C.A., Sipper, M.: Designing Breast Cancer Diagnostic via a Hybrid Fuzzy-Genetic Methodology. In: Proc. of the 1999 IEEE Int. Fuzzy Systems Conf., pp. 135–139 (1999)
[12] Pattichis, C., Schizas, C., Middleton, L.: Neural Network models in EMG diagnosis. IEEE Trans. On Biomedical Engineering 42(5), 486–496 (1995)
[13] Boulougoura, M., Wadge, E., Kodogiannis, V.S., Chowdrey, H.S.: Intelligent systems for computer-assisted clinical endoscopic image analysis. In: 2nd IASTED Int. Conf. on BIOMEDICAL ENGINEERING, Innsbruck, Austria, pp. 405–408 (2004)
[14] Czogala, E., Leski, J.: Fuzzy and Neuro-fuzzy Intelligent Systems. Springer, Heidelberg (2000)
[15] Rutkowska, D.: Fuzzy Neural Networks with an application to medical diagnosis. BioCybernetics and biomedical Engineering (1-2), 71–78 (1998)
[16] Szczepaniak, P., Lisboa, P., Kacprzyk, J.: Fuzzy Systems in Medicine. Springer, Heidelberg (2000)
[17] Jang, J.S.: ANFIS: Adaptive–network based fuzzy inference systems. IEEE Trans. On Systems, Man, & Cybernetics 23(3), 665–685 (1993)
[18] Sun, R.: Robust reasoning: Integrating rule-based and similarity-based reasoning. Artificial Intelligence 75(2), 214–295 (1995)
[19] Castro, J., Delgado, M.: Fuzzy Systems with Defuzzification are Universal Approximators. IEEE Trans. on systems, Man and Cybernetics 26, 149–152
[20] Vuorimaa, P., Jukarainen, Karpanoja, E.: A neuro-fuzzy system for chemical agent detection. IEEE Trans. On Fuzzy Systems 3(4), 415–424 (1995)
[21] Nanayakkara, T., Watanabe, K., Kiguchi, K., Izumi, K.: Fuzzy Self-Adaptive RBF Neural Network Based Control of a Seven-Link Industrial Robot Manipulator. Advanced Robotics 15(1), 17–43 (2001)
[22] Tontini, G., de Queiroz, A.: RBF Fuzzy-ARTMAP: a new fuzzy neural network for robust on-line learning and identification of patterns. In: Proc. IEEE Int. Conf. on Systems, Man & Cybernetics, vol. 2, pp. 1364–1369 (1996)
[23] Jang, J.-S.R.: Neuro-fuzzy Modelling and Control. The Proc. of the IEEE (1995)
[24] Sugeno, M., Kang, G.T.: Structure identification of fuzzy model. Fuzzy Sets and systems 28, 15–33 (1988)
[25] Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modelling and control. IEEE Trans. On systems, Man, and Cybernetics 15, 116–132 (1985)
[26] Castro, J., Delgado, M.: Fuzzy Systems with Defuzzification are Universal Approximators. IEEE Trans. On systems, Man and Cybernetics 26, 149–152
[27] Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M.: Brain-computer interface technology: A review of the first international meeting. IEEE Transactions on Rehabilitation Engineering 8(2), 164–173 (2000)
[28] Tzallas, A.T., Tsipouras, M.G., Fotiadis, D.I.: Automatic seizure detection based on time-frequency analysis and artificial neural networks. Computational Intelligence and Neuroscience 7(3), 1–13 (2007)

[29] Barry, R.J., Clarke, A.R., Johnstone, S.J.: A review of electrophysiology in attention-deficit/hyperactivity disorder:1 Qualitative and quantitive electroencephalography 2. Event-related potentials. Clinical Neurophysiology 114, 171–183, 184–198 (2003)

[30] Kovalerchuk, B., Vityaev, E., Ruiz, J.F.: Consistent Knowledge Discovery in Medical; Diagnosis. IEEE Engineering in Medicine and Biology Magazine 19(4), 26–37 (2000)

[31] Romberg, M.H.: Manual of the Nervous Disease of Man, pp. 395–401. Syndenham Society, London (1853)

[32] Paulus, W.M., Straube, A., Brandt, T.: 'Visual stabilization of posture: physiological stimulus characteristics and clinical aspects'. Brain 107, 1143–1163 (1984)

[33] Gagey, P., Gentaz, R., Guillamon, J., Bizzo, G., Bodot-Braeard, C., Debruille, Baudry, C.: Normes 1985. Association Française de Posturologie, Paris (1988)

[34] Ronda, J.M., Galvañ, B., Monerris, E., Ballester, F.: Asociación entre Síntomas Clínicos y Resultados de la Posturografía Computerizada Dinámica. Acta Otorrinolaringol. Esp. 53, 252–255 (2002)

[35] Barona, R.: Interés clínico del sistema NedSVE/IBV en el diagnóstico y valoración de las alteraciones del equilibrio. Biomechanics Magazine of the Institute of Biomechanics of Valencia, IBV (February 2003)

[36] Rocchi, L., Chiari, L., Cappello, A.: Feature selection of stabilometric parameters based on principal component analysis. In: Medical & Biological Engineering & Computing 2004, vol. 42 (2004)

[37] Demura, S., Kitabayashi, T.: 'Power spectrum characteristics of body sway time series and velocity time series of the center of foot pressure during a static upright posture in preschool children'. Sport Sciences for Health 3(1), 27–32 (2008)

[38] Diener, H.C., Dichgans, J., Bacher, M., Gompf, B.: Quantification of postural sway in normals and patients with cerebellar diseases. Electroenc. and Clin. Neurophysiol. 57, 134–142 (1984)

[39] Corradini, M.L., Fioretti, S., Leo, T., Piperno, R.: Early Recognition of Postural Disorders in Multiple Sclerosis Through Movement Analysis: A Modeling study. IEEE Transactions on Biomedical Engineering 44(11) (1997)

[40] Lara, J.A., Moreno, G., Perez, A., Valente, J.P., López-Illescas, A.: Comparing posturographic time series through events detection. In: 21st IEEE International Symposium on Computer-Based Medical Systems, CBMS 2008, June 2008, pp. 293–295 (2008)

[41] Peydro, M.F., Vivas, M.J., Garrido, J.D., Barona, R.: Procedimiento de rehabilitación del control postural mediante el sistema NedSVE/IBV. Biomechanics Magazine of the Institute of Biomechanics of Valencia, IBV (2006)

[42] Dubes, R.C.: Cluster analysis and related issues. In: Chen, C.H., Pau, L.F., Wang, P.S.P. (eds.) Handbook of pattern Recognition & Computer Vision, pp. 3–32. World Scientific Publishing Co., Inc., River Edge

[43] Rutkowska, D.: Neuro-Fuzzy Architectures and Hybrid Learning. Springer, Heidelberg (2002)

[44] Jahankhani, P., Revett, K., Kodogiannis, V., Lygouras, J.: Classification Using Adaptive Fuzzy Inference Neural Network. In: Proceedings of the Twelfth IASTED International Conference Artificial Intelligence and Soft Computing (ASC 2008), Palma de Mallorca, Spain, September 1-3 (2008), ISBN 978-0-88986-756-7

[45] Jahankhani, P., Revett, K., Kodogiannis, V.: Data Mining an EEG Dataset With an Emphasis on Dimensionality Reduction. In: IEEE Symposium on Computational Intelligence and Data Mining (CIDM), April 1-5 (2007)

[46] Jahankhani, P., Revett, K., Kodogiannis, V.: EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks. In: IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing, Sofia, Bulgaria, October 3-6, pp. 120–125 (2006)

[47] Jahankhani, P.K., Revett, V.: Automatic Detection of EEG Abnormalities Using Wavelet Transforms. WSEAS Transactions on Signal Processing 1(1), 55–61 (2005) ISSN 1790-5022

[48] Lara, J.A., Jahankhani, P., Pérez, A., Valente, J.P., Kodogianniz, V.: Classification of Stabilometric Time-Series using an Adaptive Fuzzy Inference Neural Network System. In: 10th International conference, ICAISC 2010 Zakopane, Poland, June 2010, pp. 635–643 (2010), ISBN 0302-9743

Glossary of Terms and Acronyms

Adaptive: A system that can be modified during operation to meet a specified criterion.

Artificial neural network: An artificial neural network consists of a number of a very simple and highly interconnected processors, called nenurons, which are analogous to the biological neurons in the brain.

Back-propagation: A supervised learning rule for multilayer perceptrons that operates by calculating the value of the error function for a known input, then back-propagating the error from one layer to the previous one. Each neuron has its weights adjusted so that it reduces the value of the error function until a stable state is reached.

Class: A group of objects with common attributes.

Clustering: The process of dividing a heterogeneous group of objects into homogeneous subgroups.

Consequent: Action in the If part of a rule.

Data mining: Data mining is the extraction of knowledge from data. It can also be defined at the exploration and analysis of large quantities of data in order to discover meaningful patterns and rules.

Defuzzification: Finding the best crisp representation of a given fuzzy set.

Expert: A person who has deep knowledge in the form of facts and rules and string practical experience in a narrow domain.

Fuzzification: The first step in fuzzy inference; the process of mapping crisp input into degrees to which these inputs belong to the respective fuzzy sets.

Fuzzy inference: is a process of mapping from a given input to an output by using the theory of fuzzy set.

Fuzzy rules: A conditional statement in the IF x is THEN y is B, where x and y are linguistic variables, and A and B linguistic values determined by fuzzy sets.

Fuzzy system: Fuzzy systems are well suited for modelling human decision making. Important decisions are often based on human intuition, common sense and experience, rather than on the availability and precision of data.

Membership function: The mapping that associates each element in a set with its degree of membership. It can be expressed as discrete values or as continues function.

Neural network: represent a class of general-purpose tools that are success fully applied to prediction, classification and clustering problems.

Tuning: Tuning is the most laborious and tedious part in building a fuzzy system. It often involves adjusting existing fuzzy sets and fuzzy rules.